Preregistration Guide for Psychological Research

Original Creators: Michela Distefano & Allison Leeming
Revised by: Elora Wales, Megan Kelly, and Sophia Tran

TABLE OF CONTENTS

OVERVIEW

Pre-registration is the process of creating a plan for your research and sharing it in a registry available to the public before carrying out the research. Preregistration allows the scientific community and the general public to view the objectives and progress of your research, collectively increasing the transparency, quality, reliability, and replicability of the research.

Preregistration also functions as an exercise to think deeply about the key questions of importance, and the kind of data needed to be compatible with your planned data analysis. Therefore, not only is preregistration important for open science, but it can be incredibly helpful to one's individual process at the start of every project and a valuable resource when looking back at past projects!

Pre-registration is meant to strengthen and facilitate the research process and should not be daunting enough to keep young researchers from engaging in the practice. For this reason, we've put together the following guide on how to pre-register your research with the Open Science Framework.

REGISTRATION

1. New Registrations – creation of a new registration will create a draft registration.
2. Draft Registrations – until they are submitted and approved by authors (or 48 hours have elapsed since submission), registrations will appear as draft registrations.
3. Registrations – extant registrations.

Note on Templates – The current guide is focused on how to fill out the registration when using the OSF preregistration template. Much of the general information presented here will be similar or the same to other templates should one select a different template upon creating a new registration draft.
Note on Licensing – According to OSF (Nov, 2023): A license tells others how they can use your work in the future and only applies to the information and files submitted with the registration. For more information, see this article on licenses: https://help.osf.io/article/148-licensing

PROJECT INFORMATION

Authors

In this section, you will include all the authors that are involved in the project. *Note*: It is completely fine for authors to be added onto the project later or after the pre-registration process is complete.

There are 3 levels of author permissions. Ensure you provide each author with the correct level of permission for them, otherwise they may not be able to view or edit the document:
Read: Collaborator can view the contents of the project.
Read + Write: Collaborator can view and edit the contents of the project.
Administrator: The creator of the project. The administrator has complete access to viewing and editing the project and can add or remove collaborators.

Layman Study Description

The layman study description should include the following:

- Title of study, funding, and lead researchers.
- Brief description, including the purpose and objectives of the current project. Research question(s) and hypotheses can also be included.
- Situate the current project in the scholarly literature and provide a rationale for the study.
- Briefly describe methods and procedures that will be used to obtain the data and answer the research question(s).
- References: Include author, year, title, journal, page numbers

Note: The description should be no longer than the length of an abstract. It can give some context for the proposed study, but great detail is not needed here.

Rationale and Background

This section should contain all relevant background information necessary to provide a rationale for your research question(s) and justify their investigation on a larger scale.
As such, this section should consist of a full but concise literature review of the domain in which the study is proposed, approximately presented in the format of the introduction section of a publishable journal article including full referencing. Additionally, authors are required to provide a rationale for all research questions, justification of methodology, and justification for the use of resources (financial and otherwise) in pursuit of information on the topic. This section must contain enough information to educate the uninformed reader and should emphasize information contributing to the research hypotheses.

Aims, Objectives, and Hypotheses

In this section, provide a general overview of the purpose of the project, as well as your main research questions and hypotheses.

STUDY DESIGN

Choosing the design plan

In this section, describe the overall design of your study. Your research plan should be designed to register a single study through OSF. If you have multiple experimental designs, complete a separate preregistration through OSF.

It is important to establish a specific and complete design plan, because this will allow you to produce accurate conclusions and contribute to the systematic observation of your research question. Choosing your design plan will give you an outline for the way you should make your observations. Additionally, it will help you determine the analytical and statistical procedures necessary for analyzing your data.

Examples of research designs include observational studies, experimental studies, meta-analyses, and systematic reviews.

Note: Based on your chosen study design, give all necessary details pertaining to the specifics of your design. Be sure to determine if every parameter has been specified in the description of the study design.

Variables and primary outcomes

Here, describe all variables (both manipulated and measured) that will be used in the data analysis plan. In the data analysis plan, you will be able to describe how each variable will be used to meet your objectives. If certain variables are being measured for the purpose of exploratory analyses, we encourage you to list them in the interest of transparency.

Manipulated Variables

- If applicable in the context of your research, describe variables that you plan to manipulate and the levels or treatment arms for each of these variables.
- Be sure to give a clear and specific definition of each manipulated variable. This includes a precise description of the levels at which each variable will be set. E.g., instead of describing a variable as "loud" or "quiet", give a precise decibel level or a means of reproducing each level. This is important for future researchers who aim to replicate the precise research study.
- *Note*: This is NOT applicable in observational studies.

Measured Variables

- Each variable you will measure should be described in detail here. Descriptions should be as specific as possible.
    - E.g., Instead of "intelligence" as a variable, you could say "IQ as measured by the Wechsler Adult Intelligence Scale".
- Include predictors, covariates, and outcome measures (see Primary Outcome Measures below).
- While it is not necessary to include variables that you plan on collecting if they are not going to be included in the confirmatory analyses of this specific study, we encourage you to do so for the purpose of transparency.
- *Note:* For observational studies and meta-analyses, include measured variables only.

Primary Outcome Measures

- The primary outcome measures (i.e., the dependent variable) should be the outcomes that you consider the most important to investigate in your study among many others. Primary outcome measures are in line with the researcher's specific study objectives.

Indices

- If you plan to combine measures into an index or a mean, for example, describe the specific measures you will use to do so, and how they will be combined.

- If you are using multiple parts of the data to create a single variable, explain how this will occur (i.e., data that will be included and formula or weights for each measure).
- Here, you should include a formula or a precise description of your method. If you plan to use a complex statistical method to combine your measures (e.g., factor analysis), note this here, but describe the procedure in detail in your data analysis plan.
- *Note:* For standard summary statistics (e.g., means), a formula is not required.

PARTICIPATION

Here, researchers should describe the recruitment method, the intended number of participants in the project, and the rationale for the intended number of participants. *Note:* The data described in this section should be the actual data used for analysis. If a subset of a larger dataset is being used, please describe the specific subset that will be used in this project.

Recruitment Method

Please describe the process by which you will collect your data. If you are using human subjects, this should include the following information:

- The population from which you obtain subjects
- Recruitment efforts (e.g., how will participants be recruited and who will be in charge of recruiting and contacting participants)
- How subjects will be selected for eligibility from the initial pool (e.g. inclusion and exclusion criteria)
- Compensation for participation (e.g., details about the type and amount of the compensation; will participants be compensated if they withdraw from the project?)
- The timeline of your study

For studies that do not include human subjects, include information about how you will collect samples, duration of data gathering efforts, source or location of samples, or batch numbers you will use.

Note: Information pertaining to the informed consent process should be included in the section entitled "Participant Informed Consent".

Demographics

Include any and all demographic factors pertaining to your study sample. These can include, among others, the following factors:

- Age
- Gender
- Income
- Education
- Sexual orientation
- Family life cycle
- Religion

- Socio-economic status

Note: For studies utilizing conditions requiring different demographic factors, please specify here (e.g., lower SES, higher SES).

Specific Protocol

- Describe in detail what participants will be asked to do for the study
    - *Note:* specific measures should be described in the section entitled "Participant Questionnaires, Screening, and Interviews"
- Include how often participants will be asked to participate and how long each research activity will last.
- Describe when data collection will occur (e.g., interviews with school children will take place outside of class time). Location of data collection can also be included (e.g. name of Laboratory).
- Information about whether scientific equipment involving direct or indirect physical contact will be used can be included in this section (e.g., electrodes, sensory devices, probes). If audio/video recordings will be used in the project, describe here.

Sample Size

- Describe the sample size of your study:
    - How many subjects will be analyzed in the study?
    - If applicable, how many subjects are you collecting at each level of the analysis?
    - Do you foresee any obstacles in obtaining sample size that may extend funding?
- For some studies, this will simply be the number of samples or the number of clusters. For others, this could be an expected range, minimum, or maximum number.
- *Note:* Rationale for the sample size, including all power calculations, should be detailed in the section entitled "Sample Size Power Calculations".

Stopping Rule

- If your data collection procedures do not give you full control over your exact sample size, specify how you will decide when to terminate your data collection.
    - E.g., "We will post participant sign-up slots weekly each Monday morning, with 30 slots per week. This will be repeated each week until we reach our intended number of participants."
- Additionally, if the number of interested subjects who volunteer for the study exceeds the intended number of participants for the project, please specify whether or not these individuals will be included in the study.
- If possible, it would be helpful to include a power analysis wherein one outlines the desired N for the study based on the power they are using to be able to detect a certain magnitude of effect based on the most important analysis (or analysis of most interest).
    - E.g., if performing a 2 x 2 within-participants ANOVA: "The sample size will be N = 120. This is the N required to have ~85% power to detect a Cohen's **f** of 0.28

(partial eta squared of .07) and ~over 99% power to detect a Cohen's **f** of .57 (α = .05, two-tailed, η_p^2= 0.24) in a 2w*2w design using ANOVA-exact (Lakens & Caldwell, 2021)." Also see section entitled "Sample Size Power Calculations".

Reference:
Lakens D, Caldwell AR. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science, 4*(1). doi:10.1177/2515245920951503

MATERIALS

Advertising Materials

In this section, please include any and all advertisement posters for human participants for your study.

Details to include:

- Where will these advertisements be located?
- Target populations and groups of interest. Are there separate advertisements for each group of interest?
- Is permission required for placement of certain advertisements? Please indicate whether permission has been obtained.
- What exclusion criteria apply to target groups? Are those criteria indicated on the advertisements? Why or why not.
- Will participants be financially compensated for their involvement in the study?
- Duration of participant involvement. Will they be required for more than one session?
- What information are potential participants given about the purpose of the study?

Examples of all advertisement material (online, posters, transcripts of radio announcements, etc.) should be appended.

Participant Questionnaires, Screening, and Interviews

In this section, describe and append all relevant clinical questionnaires, screening documents, and information gathering tools used in the study and to be filled out by participants. This includes but is not limited to psychometrically validated measures, screening questionnaires, demographic questionnaires, questionnaires requiring personal information, measures delivered both online and in person, and any cognitive, emotional, or skill measures/screening tools. Additionally, be sure to include transcripts of any clinical interviews. In the case of an unstructured or semi-structured interview, append topics to be covered, information targets, and goals/purpose of the interview.

Participant Informed Consent and Debriefing

In this section, please include any and all documents related to participant consent. This includes appending all relevant documentation (consent to be contacted for research, consent for imaging,

consent for invasive procedures, consent for collection of biological samples, release of personal information, etc.) from all involved parties, organizations, and institutions.

The participant informed consent form should include the following information:

- Affiliation
- Title of the study
- Principal and Co-Investigators
- Brief description of the study
- Purpose of the study
- Procedure
- Risks and benefits of participation
- Compensation for participation
- Confidentiality (i.e., measures taken to ensure confidentiality and limits of confidentiality)
- Contact information (i.e., principal investigator and ethics board)
- Name, signature and date of both participant and person obtaining consent

Two copies of the informed consent form are necessary: one for the participant and one for the research team.

Note: Keep in mind that the informed consent form is intended for participants and not for researchers. Thus, the language used in this form should be appropriate for the general public's understanding. If participants are being deceived during the study, a debriefing form detailing the reasoning behind the procedures and the necessity of concealing this information is required. Elements that should be included in a debriefing form include:

- Institution name and title of study
- Contact information for the principal investigator and ethics board
- An explanation of the specific purpose of the study and the tasks that were used during the study
- An explanation of how and why participants were deceived during the study
- A list of resources for the participant, if applicable

DATA ANALYSIS

Sample Size Power Calculations

Sample Size Rationale

- In this section, please include all rationale for sample size in your study. This should, if possible, include G*Power analyses to statistically justify sample size and expected power. These analyses should be based on the a priori statistical tests planned for the data analysis portion of the study. If this is not possible or no statistical tests are planned a priori, rationale may consist of other constraints such as time, money, or personnel availability.

- E.g., "We used the software program G*Power to conduct a power analysis. Our goal was to obtain .95 power to detect a medium effect size of .25 at the standard .05 alpha error probability. This yielded a required sample size of X."

Useful References and Tutorials for G*Power:

Buchner, A., Erdfelder, E., Faul, F., Lang, A-G. (Last updated March 1, 2017). G*Power manual. Heinrich-Heine-Universität Düsseldorf. http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPowerManual.pdf

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Susanne, Mayr & Erdfelder, Edgar & Buchner, Axel & Faul, Franz. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology. 3*. 10.20982/tqmp.03.2.p051.

Choosing the Right Statistical Test

Choosing (or not choosing) the right statistical methods for your study will be both a result of and dictate several of its important elements. These include study design, research methodology, capacity for parametric and causal inferences, quantitative vs qualitative data analysis, implications for other research in the field, and capacity to describe and relate your variables of interest.

There are several categories of statistical tests and even more individual tests. The following are some basic categories to help guide your decision-making process: one-sample tests, correlations, causal and predictive relationships, group differences, and reliability testing.

Useful References for choosing the right statistical test based on variables in your study:

Choosing the correct statistical test in SAS, STATA, SPSS and R (adapted from Choosing the Correct Statistic developed by James D. Leeper, Ph.D.) UCLA: Statistical Consulting Group. https://stats.idre.ucla.edu/other/mult-pkg/whatstat/
McDonald, J. H. (2009). Handbook of biological statistics (Vol. 2, pp. 173-181). Baltimore, MD: sparky house publishing.

Nayak, B. K., & Hazra, A. (2011). How to choose the right statistical test?. *Indian journal of ophthalmology, 59*(2), 85–86. doi:10.4103/0301-4738.77005

Data Analysis Plan

The data analysis section of a preregistration is one of its most important elements in terms of the replicability and transparency of a study. Here, it is of the utmost importance that the authors are as detailed as possible, that they are open and honest about manipulations of their data, and that they diligently track any changes to their analyses over the course of the research project. In doing so, authors make possible and encourage collaboration between research teams, honesty, transparency, and replication of their study to confirm results.

Methods for data analysis will be largely dictated by experimental design, research materials, variable types, research sample, qualitative vs quantitative data, goals of analysis, and other factors. However, there are some common elements in most data analysis that are imperative to a comprehensive and transparent preregistration. Several of these elements are listed below.

Existing Data

The distinction between existing and novel data is extremely important in the context of a preregistration. This dictates whether tests are exploratory or confirmatory, which has implications for interpretation of results. Though it is always encouraged to preregister your study whether you have seen your data or not, it is ideal to do so prior to any data analysis and especially any data "peaking". This reinforces the idea of honesty and anti-bias as mantras throughout the entire process of a research project, from conception to publication.

- There are several levels or types of "existing" data. Below are some of the circumstances in which authors may find themselves upon preregistration or data analysis:
    - Data have not been collected, seen, analyzed. Preregistration contains no information related to interpretation of prior data.
    - Data has been collected, but not observed; no "peaking" or analysis.
    - Data has been collected and accessed, but not analyzed in the context of research plan. This includes no descriptive statistics.
    - Data has been collected, accessed, and analyzed. This includes any preliminary exploratory analyses, "peaking", and generation of descriptive statistics.
- If any existing data will be used in your study, there are steps that must be taken to maintain transparency. Please describe in detail the exact extent to which data has been accessed or analyzed and acknowledge whether this access has yielded any results which may speak to the validity of hypotheses. In the case in which existing data either supports or opposes hypotheses, authors must acknowledge that these hypotheses may no longer be a priori. This conceptual shift has implications for interpretation and transparency.

Statistical Analyses

In this section, please describe in detail (tables may be used) any statistical analyses with which you are treating your data. This includes exploratory and confirmatory analyses. Additional exploratory analyses that were not pre-planned are, of course, permitted, but it is important to list all additional unplanned analyses as such. Ensure clarity in identifying independent and

dependent variables, as well as contextualizing analyses in terms of you're a priori hypotheses stated earlier.

- Important elements in this section include:
    - Statistical models and specifications of each model. This includes confirmatory and exploratory, clearly differentiated with confirmatory first. Please indicate which hypotheses are being confirmed or opposed for each analysis.
    - Specify the exact method for the analysis of interactions, subgroup analyses, additional or complex contrasts, post-hoc tests, controls, variable transformations or recoding, indices, effect size measures, and other analyses.
    - Additionally, provide criteria and adequate rationale for data exclusion, cut-off values, inference criteria, procedure for incomplete or missing data, and development of clinically meaningful thresholds.
- The most important element in this section is the aspect of *replicability*. Please ensure that analytical procedures are clear enough to easily translate across research laboratories (within reason) and leave nothing out. This caveat speaks to the overall purpose of the entire preregistration process in terms of facilitating inter-lab communication and study replication.

Bear in mind that specific requirements will change based on study type, materials and methods, and research domain. For example: neuroimaging studies will require inclusion of additional information regarding imaging parameters, pre- and post-processing procedures, normalization, and other mechanical and programming specifications. Individual demands are often study-specific, so we recommend being even more inclusive than may seem necessary. As a rule of thumb, authors should always be asking themselves: "Would another research team be able to accurately replicate our study using only information from the preregistration?".

REFERENCES

Buchner, A., Erdfelder, E., Faul, F., Lang, A-G. (Last updated March 1, 2017). G*Power manual. Heinrich-Heine-Universität Düsseldorf. http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPowerManual.pdf

Choosing the correct statistical test in SAS, STATA, SPSS and R (adapted from Choosing the Correct Statistic developed by James D. Leeper, Ph.D.) UCLA: Statistical Consulting Group. https://stats.idre.ucla.edu/other/mult-pkg/whatstat/

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Lakens D, Caldwell AR. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science, 4*(1). doi:10.1177/2515245920951503

McDonald, J. H. (2009). Handbook of biological statistics (Vol. 2, pp. 173-181). Baltimore, MD: sparky house publishing.

Nayak, B. K., & Hazra, A. (2011). How to choose the right statistical test?. *Indian journal of ophthalmology, 59*(2), 85–86. doi:10.4103/0301-4738.77005

Susanne, Mayr & Erdfelder, Edgar & Buchner, Axel & Faul, Franz. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology. 3*. 10.20982/tqmp.03.2.p051.